BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

specification with the following rewritten paragraph:

--Figures 41A and 41B show the ESI-CID-MS of a 27-mer RNA/DNA hybrid in the presence and absence of paromomycin, respectively.--

E2

Please replace the paragraph referring to Figure 42 beginning at page 26, line 3 of the specification with the following rewritten paragraph:

--Figures 42A and 42B show the ESI-MS of a 27-mer RNA/DNA hybrid target in the presence of paromomycin alone (panel a), and in the presence of both paromomycin and a combinatorial library (panel b), respectively.--

E3

Delete paragraph referring to Figure 55 beginning at page 26, line 23 of the specification.

Delete paragraph referring to Figure 56 beginning at page 26, line 25 of the specification.

Please replace the paragraph referring to Figure 57 beginning at page 26, line 27 of the specification with the following rewritten paragraph:

--Figure 55 shows a representative flow scheme showing preferred steps for a preferred SEALS strategy.--

E4

Delete paragraph referring to Figure 58 beginning at page 26, line 29 of the specification.

Delete paragraph referring to Figure 59 beginning at page 27, line 1 of the specification.

Delete paragraph referring to Figure 60 beginning at page 27, line 3 of the specification.

Delete paragraph referring to Figure 61 beginning at page 27, line 5 of the specification.

Delete paragraph referring to Figure 62 beginning at page 27, line 7 of the specification.

Delete paragraph referring to Figure 63 beginning at page 27, line 8 of the specification.

Please replace the paragraph referring to Figure 64 beginning at page 27, line 9 of the specification with the following rewritten paragraph:

--Figure 56 shows a representative flow scheme showing preferred steps for a preferred

E5

Delete paragraph referring to Figure 65 beginning at page 27, line 11 of the specification.
Delete paragraph referring to Figure 66 beginning at page 27, line 12 of the specification.
Delete paragraph referring to Figure 67 beginning at page 27, line 13 of the specification.
Delete paragraph referring to Figure 68 beginning at page 27, line 14 of the specification.
Delete paragraph referring to Figure 69 beginning at page 27, line 15 of the specification.
Delete paragraph referring to Figure 70 beginning at page 27, line 17 of the specification.
Delete paragraph referring to Figure 71 beginning at page 27, line 18 of the specification.
Delete paragraph referring to Figure 72 beginning at page 27, line 19 of the specification.
Delete paragraph referring to Figure 73 beginning at page 27, line 20 of the specification.
Delete paragraph referring to Figure 74 beginning at page 27, line 21 of the specification.
Delete paragraph referring to Figure 75 beginning at page 27, line 22 of the specification.
Delete paragraph referring to Figure 76 beginning at page 27, line 23 of the specification.
Delete paragraph referring to Figure 77 beginning at page 27, line 24 of the specification.
Delete paragraph referring to Figure 78 beginning at page 27, line 25 of the specification.
Delete paragraph referring to Figure 79 beginning at page 27, line 26 of the specification.
Delete paragraph referring to Figure 80 beginning at page 27, line 28 of the specification.
Delete paragraph referring to Figure 81 beginning at page 27, line 29 of the specification.
Delete paragraph referring to Figure 82 beginning at page 27, line 30 of the specification.
Delete pages 28-29 of the specification.
Delete paragraph referring to Figure 114 beginning at page 30, line 1 of the specification.
Delete paragraph referring to Figure 115 beginning at page 30, line 3 of the specification.
Delete paragraph referring to Figure 116 beginning at page 30, line 5 of the specification.
Delete paragraph referring to Figure 117 beginning at page 30, line 6 of the specification.
Delete paragraph referring to Figure 118 beginning at page 30, line 7 of the specification.
Delete paragraph referring to Figure 119 beginning at page 30, line 8 of the specification.
Delete paragraph referring to Figure 120 beginning at page 30, line 9 of the specification.

Delete paragraph referring to Figure 121 beginning at page 30, line 10 of the specification.

Delete paragraph referring to Figure 122 beginning at page 30, line 11 of the specification.

Delete paragraph referring to Figure 123 beginning at page 30, line 12 of the specification.

Delete paragraph referring to Figure 124 beginning at page 30, line 13 of the specification.

Delete paragraph referring to Figure 125 beginning at page 30, line 14 of the specification.

Please replace the paragraph referring to Figure 126 beginning at page 30, line 15 of the specification with the following rewritten paragraph:

--Figure 57 shows a representative Dome structure view of region 3 of IL-2 3'UTR. SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, and SEQ ID NO:51 are shown top to bottom, respectively.--

E6

Please replace the paragraph referring to Figure 127 beginning at page 30, line 16 of the specification as follows:

--Figure 58 shows a representative structure drawing of region 3 of IL-2 3'UTR.--

E7

Delete paragraph referring to Figure 128 beginning at page 30, line 17 of the specification.

Delete paragraph referring to Figure 129 beginning at page 30, line 18 of the specification.

Delete paragraph referring to Figure 130 beginning at page 30, line 19 of the specification.

Delete paragraph referring to Figure 131 beginning at page 30, line 20 of the specification.

Delete paragraph referring to Figure 132 beginning at page 30, line 21 of the specification.

Delete paragraph referring to Figure 133 beginning at page 30, line 22 of the specification.

Delete paragraph referring to Figure 134 beginning at page 30, line 23 of the specification.

Delete paragraph referring to Figure 135 beginning at page 30, line 24 of the specification.

Delete paragraph referring to Figure 136 beginning at page 30, line 25 of the specification.

Delete paragraph referring to Figure 137 beginning at page 30, line 26 of the specification.

Delete paragraph referring to Figure 138 beginning at page 31, line 1 of the specification.

Please replace the paragraph referring to Figure 139 beginning at page 31, line 4 of the specification with the following rewritten paragraph:

--Figure 59 depicts FTMS spectrum obtained from a mixture of a 16S RNA model (10 μM) and a 60-member combinatorial library.--

E8

Please replace the paragraph referring to Figure 140 beginning at page 31, line 6 of the specification with the following rewritten paragraph:

--Figure 60 depicts an expanded view of the 1863 complex from Figure 59.--

E9

Please replace the paragraph referring to Figure 141 beginning at page 31, line 7 of the specification with the following rewritten paragraph:

--Figure 61 depicts mass of a binding ligand determined from a starting library of compounds.--

E10

Please replace the paragraph referring to Figure 142 beginning at page 31, line 9 of the specification with the following rewritten paragraph:

--Figure 62 depicts high resolution ESI-FTICR spectrum of the library used in Figures 60 and 61.--

E11

Please replace the paragraph referring to Figure 143 beginning at page 31, line 11 of the specification with the following rewritten paragraph:

--Figure 63 depicts use of exact mass measurements and elemental constraints to determine the elemental composition of an exemplary "unknown" binding ligand.--

E12

Please replace the paragraph referring to Figure 144 beginning at page 31, line 13 of the specification with the following rewritten paragraph:

--Figure 64 depicts ESI-MS measurements of a solution containing a fixed concentration of RNA at different concentrations of ligand.--

E13

Delete paragraph referring to Figure 145 beginning at page 31, line 15 of the specification.

Please replace the paragraph referring to Figure 146 beginning at page 31, line 17 of the specification with the following rewritten paragraph:

E14 --Figure 65 depicts MASS screening of a 27 member library against a 27-mer RNA construct representing the prokaryotic 16S A-site.--

Please replace the paragraph referring to Figure 147 beginning at page 31, line 19 of the specification with the following rewritten paragraph:

E15 --Figure 66 depicts MS/MS of a 27-mer RNA construct representing the prokaryotic 16S A-site containing deoxyadenosine residues at the paromomycin binding site.--

Please replace the paragraph referring to Figure 148 beginning at page 31, line 21 of the specification with the following rewritten paragraph:

E16 --Figure 67 depicts MS-MS spectra obtained from a mixture of a 27-mer RNA construct representing the prokaryotic 16S A-site containing deoxyadenosine residues at the paromomycin binding and the 216 member combinatorial library respectively.--

Please replace the paragraph referring to Figure 149 beginning at page 31, line 24 of the specification with the following rewritten paragraph:

E17 --Figure 68 depicts secondary structures of the 27 base RNA models used in this work corresponding to the *18S* (eukaryotic) SEQ ID NO:382 and *16S* (prokaryotic) SEQ ID NO:35 A-sites.--

Please replace the paragraph referring to Figure 150 beginning at page 31, line 26 of the specification with the following rewritten paragraph:

E18 --Figure 69 depicts ESI-FTICR spectrum of a mixture of 27-base representations of the *16S* A-site with (7 μM) and without (1 μM) an 18 atom neutral mass tag attached to the 5- terminus in

the presence of 500 nM paromomycin.--

Please replace the paragraph referring to Figure 151 beginning at page 31, line 29 of the specification with the following rewritten paragraph:

--Figure 70 depicts mass spectra from simultaneous screening of 16S A-site and 18S A-site Model RNAs against a mixture of aminoglycosides.--

Please replace the paragraph referring to Figure 152 beginning at page 31, line 31 of the specification with the following rewritten paragraph:

--Figure 71 depicts sequences and structures for oligonucleotides R and C.--

Please replace the paragraph referring to Figures 153A, B and C beginning at page 32, line 1 of the specification with the following rewritten paragraph:

--Figure 72A depicts mass spectrum obtained from a mixture of 5 μM C and 125 nM paromomycin. Figure 72B depicts MS-MS spectrum obtained following isolation of $[M-5H]^{5-}$ ions (*m/z* 1783.6) from uncomplexed C. Figure 72C depicts MS-MS spectrum obtained following isolation of $[M-5H]^{5-}$ ions (*m/z* 1907.5) from C complexed with paromomycin.--

Please replace the paragraph referring to Figures 154A and B beginning at page 32, line 6 of the specification with the following rewritten paragraph:

--Figure 73A depicts MS-MS spectrum obtained from a mixture of 10 μM C and a 216 member combinatorial library following isolation of $[M-5H]^{5-}$ ions (*m/z* 1919.0) from C complexed with ligands of mass 676.0 ± 0.6. Figure 73B depicts MS-MS spectrum obtained from a mixture of 10 μM C and a 216 member combinatorial library following isolation of $[M-5H]^{5-}$ ions (*m/z* 1934.3) from C complexed with ligands of mass 753.5 ± 0.6.--

Please replace the paragraph referring to Figure 155 beginning at page 32, line 11 of the specification with the following rewritten paragraph:

--Figure 74 depicts electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry of a target / putative ligand mixture.--

E23

Please replace the paragraph referring to Figure 156 beginning at page 32, line 13 of the specification with the following rewritten paragraph:

--Figure 75 shows isotope clusters from the spectrum of Figure 74.--

E24

Please replace the paragraph referring to Figure 157 beginning at page 32, line 14 of the specification with the following rewritten paragraph:

--Figure 76 depicts data tabulated and stored in a relational database.--

E25

Please replace the paragraph referring to Figure 158 beginning at page 32, line 15 of the specification with the following rewritten paragraph:

--Figure 77 shows an exemplary flow chart for a computer program for effecting certain methods in accordance with the invention.--

E26

Please replace the paragraph beginning at page 38, line 11 of the specification with the following rewritten paragraph:

--Additional nucleic acid targets may be determined independently or can be selected from publicly available prokaryotic and eukaryotic genetic databases known to those skilled in the art. Preferred databases include, for example, Online Mendelian Inheritance in Man (OMIM), the Cancer Genome Anatomy Project (CGAP), GenBank, EMBL, PIR, SWISS-PROT, and the like. OMIM, which is a database of genetic mutations associated with disease, was developed, in part, for the National Center for Biotechnology Information (NCBI). OMIM is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/Omim/. CGAP, which is an interdisciplinary program to establish the information and technological tools required to decipher the molecular anatomy of a cancer cell. CGAP is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/ncicgap/. Some of these databases may contain complete

E27

E27 cont'd

or partial nucleotide sequences. In addition, nucleic acid targets can also be selected from private genetic databases. Alternatively, nucleic acid targets can be selected from available publications or can be determined especially for use in connection with the present invention.--

Please replace the paragraph beginning at page 38, line 25 of the specification with the following rewritten paragraph:

--After a nucleic acid target is selected or provided, the nucleotide sequence of the nucleic acid target is determined and then compared to the nucleotide sequences of a plurality of nucleic acids from different taxonomic species. In one embodiment of the invention, the nucleotide sequence of the nucleic acid target is determined by scanning at least one genetic database or is identified in available publications. Preferred databases known and available to those skilled in the art include, for example, the Expressed Gene Anatomy Database (EGAD) and Unigene-Homo Sapiens database (Unigene), GenBank, and the like. EGAD contains a non-redundant set of human transcript (HT) sequences and is publicly available through the Internet at the world wide web at, for example, tigr.org/tdb/egad/egad.html. Unigene is a system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters. Each Unigene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location.--

E28

Please replace the paragraph beginning at page 39, line 9 of the specification with the following rewritten paragraph:

--In addition, Unigene contains hundreds of thousands of novel expressed sequence tag (EST) sequences. Unigene is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/UniGene/. These databases can be used in connection with searching programs such as, for example, Entrez, which is known and available to those skilled in the art, and the like. Entrez is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/Entrez. Preferably, the most complete nucleic acid sequence representation available from various databases is used. The GenBank database, which is known and available to

E29

those skilled in the art, can also be used to obtain the most complete nucleotide sequence. GenBank is the NIH genetic sequence database and is an annotated collection of all publicly available DNA sequences. GenBank is described in, for example, Nuc. Acids Res., 1998, 26, 1-7, which is incorporated herein by reference in its entirety, and can be accessed by those skilled in the art through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/Web/Genbank/index.html. Alternatively, partial nucleotide sequences of nucleic acid targets can be used when a complete nucleotide sequence is not available.

E 29 cont'd

Please replace the paragraph beginning at page 39, line 24 of the specification with the following rewritten paragraph:

In another embodiment of the present invention, the nucleotide sequence of the nucleic acid target is determined by assembling a plurality of overlapping expressed sequence tags (ESTs). The EST database (dbEST), which is known and available to those skilled in the art, comprises approximately one million different human mRNA sequences comprising from about 500 to 1000 nucleotides, and various numbers of ESTs from a number of different organisms. dbEST is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/dbEST/index.html. These sequences are derived from a cloning strategy that uses cDNA expression clones for genome sequencing. ESTs have applications in the discovery of new genes, mapping of genomes, and identification of coding regions in genomic sequences. Another important feature of EST sequence information that is becoming rapidly available is tissue-specific gene expression data. This can be extremely useful in targeting selective gene(s) for therapeutic intervention. Since EST sequences are relatively short, they must be assembled in order to provide a complete sequence. Because every available clone is sequenced, it results in a number of overlapping regions being reported in the database.

E 30

Please replace the paragraph beginning at page 40, line 7 of the specification with the following rewritten paragraph:

Assembly of overlapping ESTs extended along both the 5' and 3' directions results in a

E 31

full-length "virtual transcript." The resultant virtual transcript may represent an already characterized nucleic acid or may be a novel nucleic acid with no known biological function. The Institute for Genomic Research (TIGR) Human Genome Index (HGI) database, which is known and available to those skilled in the art, contains a list of human transcripts. TIGR is publicly available through the Internet at the world wide web at, for example, tigr.org/. The transcripts were generated in this manner using TIGR-Assembler, an engine to build virtual transcripts and which is known and available to those skilled in the art. TIGR-Assembler is a tool for assembling large sets of overlapping sequence data such as ESTs, BACs, or small genomes, and can be used to assemble eukaryotic or prokaryotic sequences. TIGR-Assembler is described in, for example, Sutton, *et al.*, *Genome Science & Tech.*, **1995**, *1*, 9-19, which is incorporated herein by reference in its entirety, and is publicly available through the Internet via file transfer program at, for example tigr.org/pub.software/TIGRassembler. In addition, GLAXO-MRC, which is known and available to those skilled in the art, is another protocol for constructing virtual transcripts. In addition, "Find Neighbors and Assemble EST Blast" protocol, which runs on a UNIX platform, has been developed by Applicants to construct virtual transcripts. Preferred steps in the Find Neighbors and Assemble EST Blast protocol is described in the flowchart set forth in Figure 2. PHRAP is used for sequence assembly within Find Neighbors and Assemble EST Blast. PHRAP is publicly available through the Internet at, for example, chimera.biotech.washington.edu/uwgc/tools/phrap.htm. One skilled in the art can construct source code to carry out the preferred steps set forth in Figure 2.

E31 cont'd

Please replace the paragraph beginning at page 41, line 21 of the specification with the following rewritten paragraph:

Sequence similarity searches can be performed manually or by using several available computer programs known to those skilled in the art. Preferably, Blast and Smith-Waterman algorithms, which are available and known to those skilled in the art, and the like can be used. Blast is NCBI's sequence similarity search tool designed to support analysis of nucleotide and protein sequence databases. Blast is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gob/BLAST/. The GCG Package provides a local version of Blast that can be

E32

used either with public domain databases or with any locally available searchable database. GCG Package v.9.0 is a commercially available software package that contains over 100 interrelated software programs that enables analysis of sequences by editing, mapping, comparing and aligning them. Other programs included in the GCG Package include, for example, programs which facilitate RNA secondary structure predictions, nucleic acid fragment assembly, and evolutionary analysis. In addition, the most prominent genetic databases (GenBank, EMBL, PIR, and SWISS-PROT) are distributed along with the GCG Package and are fully accessible with the database searching and manipulation programs. GCG is publicly available through the Internet at the world wide web at, for example, gcg.com/. Fetch is a tool available in GCG that can get annotated GenBank records based on accession numbers and is similar to Entrez. Another sequence similarity search can be performed with GeneWorld and GeneThesaurus from Pangea. GeneWorld 2.5 is an automated, flexible, high-throughput application for analysis of polynucleotide and protein sequences. GeneWorld allows for automatic analysis and annotations of sequences. Like GCG, GeneWorld incorporates several tools for homology searching, gene finding, multiple sequence alignment, secondary structure prediction, and motif identification. GeneThesaurus 1.0tm is a sequence and annotation data subscription service providing information from multiple sources, providing a relational data model for public and local data.

E 32 cont'd

Please replace the paragraph beginning at page 42, line 22 of the specification with the following rewritten paragraph:

--Another toolkit capable of doing sequence similarity searching and data manipulation is SEALS, also from NCBI. This tool set is written in perl and C and can run on any computer platform that supports these languages. It is publicly available through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/Walker/SEALS/. This toolkit provides access to Blast2 or gapped blast. It also includes a tool called tax_collector which, in conjunction with a tool called tax_break, parses the output of Blast2 and returns the identifier of the sequence most homologous to the query sequence for each species present. Another useful tool is feature2fasta which extracts sequence

E 33

E33 cont'd

fragments from an input sequence based on the annotation. An exemplary use for this tool is to create sequence files containing the 5' untranslated region of a cDNA sequence.

Please replace the paragraph beginning at page 43, line 28 of the specification with the following rewritten paragraph:

In another embodiment of the invention, the sequences required are obtained by searching ortholog databases. One such database is Hovergen, which is a curated database of vertebrate orthologs. Ortholog sets may be exported from this database and used as is, or used as seeds for further sequence similarity searches as described above. Further searches may be desired, for example, to find invertebrate orthologs. Hovergen is publicly available through the Internet via file transfer program at, for example, pbil.univ-lyon1.fr/pub/hovergen/. A database of prokaryotic orthologs, COGS, is available and can be used interactively through the Internet at the world wide web at, for example, ncbi.nlm.nih.gov/COG/.

E34

Please replace the paragraph beginning at page 46, line 15 of the specification with the following rewritten paragraph:

In one embodiment of the invention, secondary structure analysis is performed by alignment and covariance analysis. Numerous protocols for alignment and covariance analysis are known to those skilled in the art. Preferably, alignment is performed by ClustalW, which is available and known to those skilled in the art. ClustalW is a tool for multiple sequence alignment that, although not a part of GCG, can be added as an extension of the existing GCG tool set and used with local sequences. ClustalW is publicly available through the Internet at, for example, dot.imgen.bcm.tmc.edu:9331/multialign/Options/clustalw.html. ClustalW is also described in Thompson, *et al.*, *Nuc. Acids Res.*, **1994**, *22*, 4673-4680, which is incorporated herein by reference in its entirety. These processes can be scripted to automatically use conserved UTR regions identified in earlier steps. Seqed, a UNIX command line interface available and known to those skilled in the art, allows extraction of selected local regions from a larger sequence. Multiple sequences from many different species can be clustered and aligned for further analysis.

E35

Please replace the paragraph beginning at page 47, line 9 of the specification with the following rewritten paragraph:

--Covariation is a process of using phylogenetic analysis of primary sequence information for consensus secondary structure prediction. Covariation is described in the following references, each of which is incorporated herein by reference in their entirety: Gutell, *et al.*, "Comparative Sequence Analysis Of Experiments Performed During Evolution" In Ribosomal RNA Group I Introns, Green, Ed., Austin:Landes, **1996**; Gautheret, *et al.*, *Nuc. Acids Res.*, **1997**, *25*, 1559-1564; Gautheret, *et al.*, *RNA*, **1995**, *1*, 807-814; Lodmell, *et al.*, *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 10555-10559; Gautheret, *et al.*, *J. Mol. Biol.*, **1995**, *248*, 27-43; *Gutell, Nuc. Acids Res.*, **1994**, *22*, 3502-3517; Gutell, *Nuc. Acids Res.*, **1993**, *21*, 3055-3074; *Gutell, Nuc. Acids Res.*, **1993**, *21*, 3051-3054; Woese, *Proc. Natl. Acad. Sci. USA*, **1989**, *86*, 3119-3122; and Woese, *et al.*, *Nuc. Acids Res.*, **1980**, *8*, 2275-2293. Preferably, covariance software is used for covariance analysis. Preferably, Covariation, a set of programs for the comparative analysis of RNA structure from sequence alignments, is used. Covariation uses phylogenetic analysis of primary sequence information for consensus secondary structure prediction. Covariation is publicly available through the Internet at the world wide web at, for example mbio.ncsu.edu/RNaseP/info/programs/programs.html. A complete description of a version of the program has been published (Brown, J. W. 1991 Phylogenetic analysis of RNA structure on the Macintosh computer. CABIOS7:391-393). The current version is v4.1, which can perform various types of covariation analysis from RNA sequence alignments, including standard covariation analysis, the identification of compensatory base-changes, and mutual information analysis. The program is well-documented and comes with extensive example files. It is compiled as a stand-alone program; it does not require Hypercard (although a much smaller 'stack' version is included). This program will run in any Macintosh environment running MacOS v7.1 or higher. Faster processor machines (68040 or PowerPC) is suggested for mutual information analysis or the analysis of large sequence alignments.--

Please replace the paragraph beginning at page 48, line 3 of the specification with the following rewritten paragraph:

In another embodiment of the invention, secondary structure analysis is performed by secondary structure prediction. There are a number of algorithms that predict RNA secondary structures based on thermodynamic parameters and energy calculations. Preferably, secondary structure prediction is performed using either M-fold or RNA Structure 2.52. M-fold is publicly available through the Internet at the world wide web at, for example, ibc.wustl.edu/-zuker/ma/form2.cgi or can be downloaded for local use on UNIX platforms. M-fold is also available as a part of GCG package. RNA Structure 2.52 is a windows adaptation of the M-fold algorithm and is publicly available through the Internet at, for example, 128.151.176.70/RNAstructure.html.

Please replace the paragraph beginning at page 51, line 5 of the specification with the following rewritten paragraph:

In one embodiment of the invention, nucleic acids having secondary structure which correspond to the structure descriptor elements are identified by searching at least one database. Any genetic database can be searched. Preferably, the database is a UTR database, which is a compilation of the untranslated regions in messenger RNAs. A UTR database is publicly available through the Internet via file transfer program at area.ba.cnr.it/pub/embnet/database/utr/. Preferably the database is searched using a computer program, such as, for example, Rnamot, a UNIX-based motif searching tool available from Daniel Gautheret. Each "new" sequence that has the same motif is then queried against public domain databases to identify additional sequences. Results are analyzed for recurrence of pattern in UTRs of these additional ortholog sequences, as described below, and a database of RNA secondary structures is built. One skilled in the art is familiar with Rnamot. Briefly, Rnamot takes a descriptor string, such as the one shown in Figure 9, and searches any Fasta format database for possible matches. Descriptors can be very specific, to match exact nucleotide(s), or can have built-in degeneracy. Lengths of the stem and loop can also be specified. Single stranded loop regions can have a variable length. G-U pairings are allowed and can be specified as a wobble parameter. Allowable mismatches can also be included in the descriptor definition. Functional significance is assigned to the motifs if their biological role is known based on previous analysis. Known regulatory regions such as Iron Response Element have been found using this technique (see,

E 38

Example 1 below). In embodiments of the invention in which a database containing prokaryotic molecular interaction sites is compiled, it is preferable to refrain from searching human sequences or, alternatively, discarding human sequences when found.

Please replace the paragraph beginning at page 71, line 13 of the specification with the following rewritten paragraph:

Using DOCK, ligands have been identified for certain protein targets. Recent efforts in this area have resulted in reports of the use of DOCK to identify and design small molecule ligands that exhibit binding specificity for nucleic acids such as RNA double helices. While RNA plays a significant role in many diseases such as AIDS, viral and bacterial infections, few studies have been made on small molecules capable of specific RNA binding. Compounds possessing specificity for the RNA double helix, based on the unique geometry of its deep major groove, were identified using the DOCK methodology (Chen *et al., Biochemistry*, **1997**, *36*, 11402; *Kuntz et al., Acc. Chem. Res.*, **1994**, *27*, 117). Using a recent X-ray structure for r(UAAGGAGGUGAU).r(AUCACCUCCUUA) (SEQ ID NO:52) as the model structure for the A-form RNA duplex, DOCK identified several aminoglycosides as candidate ligands, characterized by shape complementarity to the RNA groove. Binding experiments then revealed that one of these aminoglycosides not only bound preferentially to RNA over B-form DNA but also that the ligand binds in the targeted RNA major groove. Recently, the application of DOCK to the problem of ligand recognition in DNA quadruplexes has also been reported (Chen *et al., Proc. Natl. Acad. Sci.*, **1996**, *93*, 2635).

E 39

Delete paragraph beginning at page 72, line 28 of the specification.

Delete pages 73-88 of the specification.

Delete paragraph beginning at page 88, line 18 of the specification.

Delete paragraph beginning at page 89, line 3 of the specification.

Please replace the paragraph beginning at page 140, line 5 of the specification with the following rewritten paragraph:

--The preferred model system employed herein comprises a library comprised of five 2-deoxystreptamine aminoglycoside antibiotics which have a range of binding affinities for the decoding sites of the prokaryotic and eukaryotic ribosomal RNA ranging from ~28 nM to ~ 1.5 mM. Figure 68 illustrates the secondary structures for the 27-nucleotide models of the 16S and 18S rRNA decoding sites. These constructs consist of a 7 base pair stem structure containing a non-canonical U-U and a purine-adenosine mismatch base pair adjacent to a bulged adenosine residue closed by a UUCG tetraloop. NMR studies of a complex between 16S and paromomycin show that the RNA makes primary hydrogen bond, electrostatic, and stacking contacts with the aminoglycoside (Fourmy, *et al., Science*, **1996**, *274*, 1367-1371) and that paromomycin binds in the major groove of the model A-site RNA within the pocket created by the A-A base pair and the single bulged adenine. The masses for the two RNA models differ by only 15.011 Da and the $(M-5H^+)^{5-}$ species of these constructs differ by only 3 m/z units. While the high resolution capabilities of the FTICR mass spectrometer can easily resolve these species, mass spectra from a solution containing both RNAs are complicated by overlap among the signals from free RNA ions and their sodium and potassium-adducted species.--

E40

Please replace the paragraph beginning at page 141, line 9 of the specification with the following rewritten paragraph:

--While the ability to shift the m/z range of closely related macromolecules is highly desirable as described above, it is preferably desired that the mass tag does not alter key physical properties of the target or the ligand binding properties. Preferably, an 18-atom mass tag ($C_{12}H_{25}O_9$) attached to the 5'-terminus of the RNA oligomer through a phosphodiester linkage can be employed. This mass tag has no appreciable affect on oligonucleotide solubility, ionization efficiency, or UV absorbance, and does not alter RNA-ligand binding. This latter attribute is evidenced by the data in Figure 69 that illustrates the conserved ratio of free:bound RNA for the untagged and tagged RNA models of the bacterial decoding site under competitive binding conditions with paromomycin.--

E41

Please replace the paragraph beginning at page 141, line 28 of the specification with the following rewritten paragraph:

--The ESI-FTICR mass spectrum depicted in Figure 70 was acquired from a 10 mM mixture of untagged 16S and tagged 18S in the presence of an equimolar mixture of five aminoglycosides. It is to be understood that other biomolecules may be used in place of the aminoglycosides. The aminoglycosides have been selected from two classes of 2-deoxystreptamines: 4,5-disubstituted (paromomycin, and lividomycin), and 4,6-disubstituted (tobramycin, sisomicin, and bekanamycin), present at 500 nM each. Complexes corresponding to 1:1 binding of individual aminoglycosides were observed between 16S and all members of the aminoglycoside mixture, with the apparent affinities estimated from the abundances of the respective complexes differing substantially. Signal intensities from the complexes with paromomycin (m/z 1925.572) and lividomycin (m/z 1954.790) are consistent with MS-measured dissociation constants of 110 nM and 28 nM, respectively. The intensities of 16S complexes with tobramycin (m/z 1895.960), bekanamycin (m/z 1899.171), and sisomicin (m/z 1891.972) were reduced, consistent with solution dissociation constants of ~1.5 mM. Wang, *et al., Biochemistry*, **1997**, *36*, 768-779. Hence, under these assay conditions, the MS-observed ion abundances reflect the solution dissociation constants. The inset in Figure 70 demonstrates the ability to resolve the isotopic envelope for each complex and allows mass differences to be calculated from homo-isotopic species, thus, measuring the difference in m/z between the RNA target and the RNA-ligand complex allows precise mass determination of the ligand. The spectrum is calibrated using multiple isotope peaks of the $(M-5H^+)^{5-}$ and $(M-4H^+)^{4-}$ charge states of the free RNA as internal mass standards which brackets the m/z range in which complexes are observed. The average mass measurement error obtained for the complexes in Figure 70 is 2.1 ppm when m/z differences are measured between the most abundant (4 ^{13}C) isotope peak of 16S and each complex. This post calibration scheme is easily automated which enables rapid, high precision mass measurements of affinity selected ligands against multiple targets in a high throughput mode.--

E42

Please replace the paragraph beginning at page 142, line 22 of the specification with the following rewritten paragraph:

--The enhanced affinity of lividomycin for 16S relative to the affinity of paromomycin for 16S is interesting. While lividomycin is believed to bind to the 16S ribosomal subunit, the exact site of interaction has not been established. Lividomycin has two significant structural differences from paromomycin. First, the additional mannopyranosyl ring could generate new macromolecular contacts with the RNA. However, the orientation of paromomycin ring IV is disordered in the NMR-derived structure for the complex with 16S. In addition, a hydroxyl group on ring I that makes a contact with A1492 is missing. The relatively high abundance of the 16S-lividomycin complex suggests that lividomycin binds at or near the 16S A-site, and generates additional contacts that enhance the binding affinity nearly 4-fold. Perhaps the most striking feature of the spectrum in Figure 70 is the complete absence of complexes between 18S and paromomycin or lividomycin. This result suggests there must be poor shape and electrostatic complementarity between the 4,5-disubstituted 2-DOS class of aminoglycoside and the conserved architecture of the eukaryotic ribosomal decoding site.--

E43

Please replace the paragraph beginning at page 144, line 3 of the specification with the following rewritten paragraph:

--In other preferred embodiments, the nucleic acid fragment comprise the consensus sequence NNNNCNNNNNNNUNNANNNNNNNN (SEQ ID NO:1) or NNNNCNNNNNNNUNNANNNNNNNN (SEQ ID NO:65) and wherein the sequence has a first double stranded region, an internal loop region, a second double stranded region and an end loop region. In other preferred embodiments, an *in silico* representation of a nucleic acid fragment that is conserved across at least two species comprises the consensus sequence NNNNCNNNNNNNUNNANNNNNNNN (SEQ ID NO:1) or NNNNCNNNNNNNUNNANNNNNNNN (SEQ ID NO:65). In other preferred embodiments, a purified and isolated nucleic acid fragment that is conserved across at least two species comprises the sequence NNNNCNNNNNNNUNNANNNNNNNN (SEQ ID NO:1) or NNNNCNNNNNNNUNNAN

E44

E4 cont'd

NNNNNNN (SEQ ID NO:65). In other preferred embodiments, a purified and isolated nucleic acid fragment comprises the human sequence (SEQ ID NO:2) UUUACAACAUAAUCUAGUUUACAG AAAAAUC. In other preferred embodiments, an *in silico* representation of a nucleic acid fragment comprises the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (SEQ ID NO:2).

Please replace the paragraph beginning at page 145, line 19 of the specification with the following rewritten paragraph:

In other preferred embodiments, the nucleic acid comprises the consensus sequence NNNNANAUGGGNNNUCACANNUANCUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:3), NNNNANAUGGGNNNUCACANNUACUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:66), NNNNANAUGGGNNNUCACANNUACUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:67), NNNNANAUGGGNNNUCACANNUACUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:68), NNNNANAUGGGNNNUCACANNUANCUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:69), NNNNANAUGGGNNNUCACANNUANCUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:70), NNNNANAUGGGNNUCACANNUANCUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:71), NNNNANAUGGGNNUCACANNUACUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:72), NNNNANAUGGGNNUCACANNUACUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:73), NNNNANAUGGGNNUCACANNUACUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:74), NNNNANAUGGGNNUCACANNUANCUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:75), NNNNANAUGGGNNUCACANNUANCUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:76), NNNNANAUGGGNUCACANNUANCUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:77), NNNNANAUGGGNUCACANNUACUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:78), NNNNANAUGGGNUCACANNUACUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:79), NNNNANAUGGGNUCACANNUACUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:80), NNNNANAUGGGNUCACANNUANCUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:81), NNNNANAUGGGNUCACANNUANCUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:82), NNNNANAUGGGUCACANNUANCUGUGUUCCUAUGGAAACUNNUUN (SEQ ID

E45

NO:83), NNNNANAUGGGUCACANNUACUGUGUUCCUAUGGAAACUNNUUN (SEQ ID NO:84), NNNNANAUGGGUCACANNUACUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:85), NNNNANAUGGGUCACANNUACUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:86), NNNNANAUGGGUCACANNUANCUGUGUUCCUAUGGAAACUNUUN (SEQ ID NO:87), or NNNNANAUGGGUCACANNUANCUGUGUUCCUAUGGAAACUUUN (SEQ ID NO:88) and having a first double stranded region, a first internal loop region, a second double stranded region, a second internal loop region, a third double stranded region and an end loop region. In other preferred embodiments, a purified and isolated nucleic acid fragment comprises the human sequence (SEQ ID NO:4) UAGGAUAUGGGUCACACUUAUCUGUGUUCCUAUGGAAACUA UUUG. In other preferred embodiments, a purified and isolated nucleic acid fragment comprises the mouse sequence (SEQ ID NO:5) UAGGAGAUGGGGGUCACACUUACUGUGUUCCUAUGGA AACUUUG. In other preferred embodiments, a purified and isolated nucleic acid fragment comprises the rat sequence (SEQ ID NO:6) UAGGAGAUGGGGGGUCACACUUACUGUGUUCC UAUGAAACUUUUG.--

Please replace the paragraph beginning at page 146, line 26 of the specification with the following rewritten paragraph:

--In other preferred embodiments, a nucleic acid comprises the consensus sequence AUGGGNNNUCACANNUANCUGUGUUCCUAU (SEQ ID NO:7), AUGGGNNNUCACANNU ACUGUGUUCCUAU (SEQ ID NO:89), AUGGGNNUCACANNUANCUGUGUUCCUAU (SEQ ID NO:90), AUGGGNNUCACANNUACUGUGUUCCUAU (SEQ ID NO:91), AUGGGNUCACA NNUANCUGUGUUCCUAU (SEQ ID NO:92), AUGGGNUCACANNUACUGUGUUCCUAU (SEQ ID NO:93), AUGGGUCACANNUANCUGUGUUCCUAU (SEQ ID NO:94), or AUGGGUCACANNUACUGUGUUCCUAU (SEQ ID NO:95) and having a first double stranded region, an internal loop region, a second double stranded region and an end loop region. A purified and isolated nucleic acid fragment comprising the human sequence (SEQ ID NO:8) AUGGGUCACACUUAUCUGUGUUCCUAU. In other preferred embodiments, a purified and isolated nucleic acid fragment comprising the mouse sequence (SEQ ID NO:9)

E46 cont'd AUGGGGGUCACACUUACUGUGUUCCUAU. In other preferred embodiments, a purified and isolated nucleic acid fragment comprising the rat sequence (SEQ ID NO:10) AUGGGGGGUCACACUUACUGUGUUCCUAU.--

Please replace the paragraph beginning at page 151, line 16 of the specification with the following rewritten paragraph:

--In other preferred embodiments, a nucleic acid comprising the consensus sequence NNUNNNNNNNGAUCNUNNNNGAUNCUUUNUNNNANNCCNNNNNNN (SEQ ID NO:20), NNUNNNNNNNGAUCNUNNNNGAUNCUUUNUNNNANCCNNNNNNN (SEQ ID NO: 96), or NNUNNNNNNNGAUCNUNNNNGAUNCUUUNUNNNACCNNNNNNN (SEQ ID NO:97) and having a first double stranded region, a first internal loop region, a second double stranded region, and a first end loop region, a third double stranded region, and a second end loop region. In other preferred embodiments, a purified and isolated nucleic acid fragment comprising the human sequence (SEQ ID NO:21) UAUAAAUAUGGAUCUUUUAUGAUUCUUUUUGUAAGCCCUA GGGGC. In other preferred embodiments, a purified and isolated nucleic acid fragment comprising the mouse or rat sequence (SEQ ID NO:22) GAUAAAUAUGGAUCUUUAAAGAUUCUUUUUG UAAGCCCCAAGGGC.--

E47

Please replace the paragraph beginning at page 152, line 9 of the specification with the following rewritten paragraph:

--In preferred embodiments, the nucleotides forming the first side of the first double stranded region are of the sequence NNNGA, UAAGA, AAAGA, UAUGA, or UUUGA and the nucleotides forming the second side of the first doubled stranded region are of the sequence UUNNG, UUUUG, or UUCUG. In other preferred embodiments, the nucleotides forming the first end loop region are of the sequence UNCU, UUCU, or UCCU. In other preferred embodiments, the nucleotides forming the first side of the second double stranded region are of the sequence AGCCC and the nucleotides forming the second side of the second doubled stranded region are of the sequence GNGNN, GGGCU, or GCGUG. In other preferred embodiments, the nucleotides forming

E48

E48 cont'd

the second end loop region are of the sequence NAN, UAC, UAG, CAA, or UAA. Preferably, the nucleic acid comprises a portion of interleukin-2 RNA. More preferably, the nucleic acid comprises a portion of the 3' UTR of interleukin-2 mRNA.

Please replace the paragraph beginning at page 159, line 4 of the specification with the following rewritten paragraph:

An alternate, and preferred, approach to finding orthologs is the use of Hovergen database and query tools that have been described in Duret, *et al., Nuc. Acids Res.*, **1994**, *22*, 2360-2365, which is incorporated herein by reference in its entirety. Hovergen was used to identify related sequences (tree classification at the species level and classification at the order level). Sequences corresponding to each of these orthologs was saved in GenBank format and grouped together in a single data file. Untranslated regions in both the 5' and 3' flanks of the coding region was extracted using SEALS and COWX, as shown in Figure 55.

E49

Please replace the paragraph beginning at page 159, line 13 of the specification with the following rewritten paragraph:

The IRE sequences are more constrained because they form an important structure. Thus, they stand out better and can be more readily identified even in closely related sequences. However, for this to work for any gene, the compare algorithm has been rewritten (see, Figures 5A-C). This new tool, CompareOverWins, allows a dynamic selection of both the range of window sizes, as well the hit threshold. This algorithm needs as its input parsed and separated 5' and 3' UTR sequences. Tools available within the Seals genome analysis package described earlier can be used to achieve this. Figure 55 describes the steps involved.

E50

Please replace the paragraph beginning at page 159, line 20 of the specification with the following rewritten paragraph:

To identify the IRE using the methods described herein, the compare over windows algorithm was used and the results visualized using AlignHits (Figure 5D for the algorithm). In

E51

E51 cont'd

addition to optimizing the thresholding, CompareOverWins also extracts the sequence corresponding to the hits. ClustalW (version 1.74) was used on the extracted sequences to create a locally gapped alignment. A representative flow scheme for this approach is shown in Figure 56.--

Please replace the paragraph beginning at page 159, line 28 of the specification with the following rewritten paragraph:

--Sets of sequences that show evidence of conservation in orthologs and paralogs or other related genes are analyzed for the ability to form internal structure. This is accomplished by analyzing each sequence in a matrix where the seqeunce is plotted 5' to 3' on the X axis and its complement is plotted 5' to 3' on the Y axis, such as in, for example, self-complementary analysis. Matches that correspond to potential intramolecular base pairs are scored according to a table of values. When the human ferritin IRE sequence is analyzed in this fashion, the diagonals indicate potential self- complementary regions. Each of the 13 IRE sequences described in this example were analyzed in the same fashion. While each of the sequences can form a variety of different structures, the structure most likely to occur is one common to all the sequences. By superimposing the plots of all 13 individual sequences, the potential structure common to all the sequences is deduced.--

E52

Please replace the paragraph beginning at page 160, line 10 of the specification with the following rewritten paragraph:

--The above scheme has been implemented algorithmically into a program called RevComp (*see*, Figure 53). RevComp creates a sorted list of all the structures. Representative results can be viewed either as a "dome" ouptut or as a "connect" or "ct" file which can be used in one of many RNA structure viewing programs (RNAStructure, RNAViz, etc.).--

E53

Please replace the paragraph beginning at page 160, line 22 of the specification with the following rewritten paragraph:

--Phylogenetic tree outputs for all Histone orthologs in Hovergen database was obtained. Each of these orthologs was saved in GenBank format and grouped together in a single data file.

E54

E54 contd

Untranslated regions in both the 5' and 3' flanks of the coding regions were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).--

Please replace the paragraph beginning at page 160, line 27 of the specification with the following rewritten paragraph:

--Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. The sequences corresponding to the region of interest was extracted from all species for alignment with CLUSTAL W (1.74). Following extraction of sequence information from Align Hits, CLUSTAL W (1.74) was used to provide multiple sequence alignment shown. Each of the putative hit sequences was analyzed for the ability to form internal structure. This was accomplished by analyzing each sequence in a matrix where the sequence was plotted 5' to 3' on the X axis and its complement is plotted 5' to 3' on the Y axis. Base-pairs along the diagonals indicate potential self-complementary regions that can form secondary structures. A representative sequence alignment in a dome format can show potential stem formation between the base pairs. Following conversion of the dome format file to a ct file, RNA Structure 3.21 is used to visualize the structure.--

E55

Please replace the paragraph beginning at page 161, line 11 of the specification with the following rewritten paragraph:

--Vimentin is an intermediate filament protein whose 3'UTR is highly conserved between species. Previous studies by Zehner *et al.*, (*Nuc. Acids Res.*, **1997**, *25*, 3362-3370) has shown that a proposed a complex stem-loop structure contained within this region may be important for vimentin mRNA functions such as mRNA localization. The same region was identified using the present analysis, thus validating the present approach. In addition, based on the analyses described herein, a second stem-loop structure that occurs downstream of the previously proposed structure that may have a role in regulating vimentin fuction as well has been identified.--

E56

Please replace the paragraph beginning at page 161, line 19 of the specification with the following rewritten paragraph:

E57

--A representative phylogenetic tree output for all Vimentin orthologs in Hovergen database was obtained. Each of these orthologs was saved in GenBank format and grouped together in a single data file. Untranslated regions in both the 5' and 3' flanks of the coding regions were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).-

Please replace the paragraph beginning at page 161, line 24 of the specification with the following rewritten paragraph:

E58

--Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. Two such regions appeared, and were used for subsequent analyses. Following extraction of sequence information from Align Hits for the first region, CLUSTAL W was used to provide multiple sequence alignment. Potential stem formation between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure. This structure is very similar to the one proposed by Zehner *et al.* Zehner *et al.* presented a detailed chemical analysis of their proposed structure for the minimal binding domain in the 3' UTR of Vimentin. This analysis included cleavage with single-strand-specific (ChS or T1) or double-strand-specific (V1) nucleases as well as after exposure to lead acetate.-

Please replace the paragraph beginning at page 162, line 6 of the specification with the following rewritten paragraph:

E59

--Following extraction of sequence information from Align Hits for the second region, CLUSTAL W was used to provide multiple sequence alignment. The potential stem formation between base pairs in the second region was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure for the second region.-

Please replace the paragraph beginning at page 162, line 19 of the specification with the following rewritten paragraph:

--A representative phylogenetic tree output for all Transferrin receptor orthologs in Hovergen database was obtained. Each of these orthologs was saved in GenBank format and grouped together in a single data file. Untranslated regions in both the 5' and 3' flanks of the coding region were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).--

E60

Please replace the paragraph beginning at page 162, line 24 of the specification with the following rewritten paragraph:

--Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. The first region, between base pairs 920 to 990, in the 3 prime UTR of transferrin receptor was extracted from all species for alignment with CLUSTAL W (1.74).--

E61

Please replace the paragraph beginning at page 162, line 29 of the specification with the following rewritten paragraph:

--Following extraction of sequence information from Align Hits for the first region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. A representative potential stem formation between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure. The second region, between base pairs 990 to 1050, in the 3 prime UTR of transferrin receptor was extracted from all species for alignment with CLUSTAL W (1.74).--

E62

Please replace the paragraph beginning at page 163, line 8 of the specification with the following rewritten paragraph:

--Following extraction of sequence information from Align Hits for the second region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Potential stem formation

E63

between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure. Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. The third region, between base pairs 1372 to 1423, in the 3 prime UTR of transferrin receptor was extracted from all species for alignment with CLUSTAL W (1.74).

E63

Please replace the paragraph beginning at page 163, line 18 of the specification with the following rewritten paragraph:

Following extraction of sequence information from Align Hits for the third region, CLUSTAL W (1.Ex.34) was used to provide multiple sequence alignment. Potential stem formation between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure. Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. The fourth region, between base pairs 1439 to 1479, in the 3 prime UTR of transferrin receptor was extracted from all species for alignment with CLUSTAL W (1.74).

E64

Please replace the paragraph beginning at page 163, line 28 of the specification with the following rewritten paragraph:

Following extraction of sequence information from Align Hits for the fourth region, CLUSTAL W (1.Ex.34) was used to provide multiple sequence alignment. Potential stem formation between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure. Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. The fifth region, between base pairs 1479 to 1542, in the 3 prime UTR of transferrin receptor was extracted from all species for alignment with CLUSTAL W (1.74).

E65

Please replace the paragraph beginning at page 164, line 7 of the specification with the following rewritten paragraph:

-Following extraction of sequence information from Align Hits for the fifth region, CLUSTAL W (1.Ex.34) was used to provide multiple sequence alignment. Potential stem formation between base pairs was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure.-

E66

Please replace the paragraph beginning at page 164, line 13 of the specification with the following rewritten paragraph:

-Orinithine decarboxylase (ODC) is the first enzyme in the polyamine biosynthetic pathway. Studies have shown existence of translational regulatory elements both in the 5' and 3' untranslated regions (Grens *et al.*, *J. Biol. Chem.*, **1990**, *265*, 11810). Secondary structures have been proposed to exist in both these regions, though there is no conclusive evidence for it. The methods described herein identified two structures in the 3'UTR, as shown below. The presence of one of these structures was verified using mass spectrometry probing (Griffey, *et al., Proc. SPIE-Int. Soc. Opt. Eng.*, 2985 (Ultrasensitive Biochemical Diagnostics II): 82-86, which is incorporated herein by reference in its entirety). Two representative sequences that showed slight variation in their lengths were made into RNA and subjected to MS structure probing. Results confirm the presence of a stem-loop structure. Accordingly, identification of a novel secondary structure can be identified from the methods described herein, and such existence has been independently verified by structure probing.-

E67

Please replace the paragraph beginning at page 164, line 26 of the specification with the following rewritten paragraph:

-Phylogenetic tree outputs for all Ornithine Decarboxylase orthologs in Hovergen database were obtained. Each of these orthologs was saved in GenBank format and grouped together in a

E68

B68 single data file. Untranslated regions in both the 5' and 3' flanks of the coding region were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).

Please replace the paragraph beginning at page 165, line 1 of the specification with the following rewritten paragraph:

Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions. Two such regions appeared, and were used for subsequent analyses. Following extraction of sequence information from the first region, CLUSTAL W (1.74) was used to provide multiple sequence alignment shown. Each of the putative hit sequences was analyzed for the ability to form internal structure in a reverse complement matrix. This was accomplished by analyzing each sequence in a matrix where the sequence is plotted 5' to 3' on the X axis and its complement is plotted 5' to 3' on the Y axis. Base-pairs along the diagonals indicate potential self- complementary regions that can form secondary structures. Domes view of the potential stem formation between base pairs in region 1 is given above the sequence alignment was determined using RevComp. RNA Structure 3.2 was used to visualize the structure.

E69

Please replace the paragraph beginning at page 165, line 13 of the specification with the following rewritten paragraph:

Mass spectrometry analyses techniques were used to probe for structure. The cluster alignment of the first region of ornithine decarboxylase 3' UTR showed presence of gaps/inserts in the multiple alignment. Two representative RNAs (gi404561 and gi35135) from the alignments were used for this experiment. Analysis of the pattern of induced fragmentation showed a very strong likelihood for base-paring along the top half of the stem-loop structure. This corresponds to bases 11-14 and 20-23 in 404561 or bases 8-11 and 18-21 in 35135. Bulged bases (G9 in 404561 or U22 in 35135) also showed characteristic fragmentation pattern. The bottom-half of the structure appeared to be less stable, and showed some fragmentation where our analyses had predicted base-paring. This was particularly true in the sequence 35135. This region, however, has several

E70

E70

contiguous A-U or G-U base-pairs which tend to be less stable, and therefore have a higher probability of fragmentation.--

Please replace the paragraph beginning at page 165, line 24 of the specification with the following rewritten paragraph:

--Following extraction of sequence information from Align Hits for the second region, CLUSTAL W was used to provide multiple sequence alignment. Potential stem formation between base pairs in the second region was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure for the second region.--

E71

Please replace the paragraph beginning at page 166, line 2 of the specification with the following rewritten paragraph:

--A representative phylogenetic tree output for all IL-2 orthologs in Hovergen database was obtained. Each of these orthologs was saved in GenBank format and grouped together in a single data file. Untranslated regions in both the 5' and 3' flanks of the coding region were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).--

E72

Please replace the paragraph beginning at page 166, line 7 of the specification with the following rewritten paragraph:

--Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions in the 3'UTR region. Two such regions appear, and were used for subsequent analyses. Following extraction of sequence information from Align Hits for the first region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Domes view of the potential stem formation between base pairs in the first region was given above the sequence alignment using RevComp. RNA Structure 3.2 was used to visualize the structure. Following extraction of sequence information from Align Hits for the second region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Potential stem formation between base pairs in the second region was given above the sequence alignment in a dome format. Following conversion of

E73

E73

the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure for the second region.

Please replace the paragraph beginning at page 166, line 20 of the specification with the following rewritten paragraph:

In addition to the two regions described above, a third region, downstream of, and partially overlapping the second region, was identified using an alternate reference sequence (3087784.fa). Following extraction of sequence information from Align Hits for this region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Potential stem formation between base pairs in the third region is shown in Figure 57 above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure for the third region (*see*, Figure 58).

E74

Please replace the paragraph beginning at page 166, line 29 of the specification with the following rewritten paragraph:

Representative phylogenetic tree output for all IL-4 orthologs in Hovergen database was obtained. Each of these orthologs was saved in GenBank format and grouped together in a single data file. Untranslated regions in both the 5' and 3' flanks of the coding region were extracted and compared using SEALS and COWX as described earlier (*see*, Figures 55 and 56).

E75

Please replace the paragraph beginning at page 167, line 4 of the specification with the following rewritten paragraph:

Following extraction and comparison by SEALS and COWX, Align Hits was used to determine potentially interesting regions in the 5'UTR region. Following extraction of sequence information from Align Hits for the above region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Domes view of the potential stem formation between base pairs in the region was given above the sequence alignment using RevComp. RNA Structure 3.2 was used to visualize the structure.

E76

Please replace the paragraph beginning at page 167, line 11 of the specification with the following rewritten paragraph:

--Align Hits was used to view hits in the 3'UTR region of IL-4. Following extraction of sequence information from Align Hits for the 3' UTR region, CLUSTAL W (1.74) was used to provide multiple sequence alignment. Potential stem formation between base pairs in the second region was given above the sequence alignment in a dome format. Following conversion of the dome format file to a ct file, RNA Structure 3.21 was used to visualize the structure for the second region.--

E77

Please replace the paragraph beginning at page 183, line 11 of the specification with the following rewritten paragraph:

--Cleavage and fragmentation of the complex by CID afforded information regarding the location of binding of the paromomycin to the chimeric nucleic acid. CID was found to produce no fragmentation at the dA sites in the nucleic acid. Thus paromomycin must bind at or near all three dA residues. Paromomycin therefore is believed to bind to the dA bulge in this RNA/DNA chimeric target, and induces a conformational change that protects all three dA residues from being cleaved during mass spectrometry. See Figures 41A and 41B.--

E78

Please replace the paragraph beginning at page 184, line 7 of the specification with the following rewritten paragraph:

--The ESI mass spectrum so obtained, shown in Figures 42A and 42B, demonstrated the presence of new signals for the (M-5H)5- ions at m/z values of 1897.8, 1891.3 and 1884.4. Comparing these new signals to the ion peak for the 27-mer alone the observed values of m/z of those members of the combinatorial library that are binding to the target can be calculated. The masses of the binding members of the library were determined to be 566.5, 534.5 and 482.5, respectively. Knowing the structure of the scaffold, and substituents used in the generation of this library, it was possible to determine what substitution pattern (combination of substituents) was present in the binding molecules.--

E79

Please replace the paragraph beginning at page 195, line 22 of the specification with the following rewritten paragraph:

--The QXP method was used to derive an accurate structure of a bound ligand to the RNA target. The NMR structure of the bacterial 16S ribosomal A site bound to paromomycin (Fourmy *et al., Science*, **1996**, *274*, 1367; PDB ID: 1pbr) was used as the reference state. The aminoglycoside antibiotic was removed from the ligand-RNA complex. The conformation space of paromomycin was exhaustively searched using the QXP method for the lowest energy conformers. The target RNA was held rigid whereas the paromomycin was treated as fully flexible. Multiple docking searches with the randomly disrupted paromomycin as initial structures were performed. The representative lowest energy structure identified from the search (dark grey) is superimposed on the NMR structure (light grey) of the bound complex.--

E80

Please replace the paragraph beginning at page 196, line 15 of the specification with the following rewritten paragraph:

--FTMS spectrum was obtained from a mixture of a 16S RNA model (10 mM) and a 60-member combinatorial library. Signals from complexes are highlighted in the insert. Binding of a combinatorial library containing 60 members to the 16S RNA model have been examined under conditions where each library member was present at 5-fold excess over the RNA. As shown in Figure 59, complexes between the 16S RNA and ~5 ligands in the library were observed.--

E81

Please replace the paragraph beginning at page 196, line 21 of the specification with the following rewritten paragraph:

--An expanded view of the 1863 complex from Figure 59 is shown in Figure 60. Two of the compounds in the library had a nominal mass of 398.1 Da. Their calculated molecular weights based on molecular formulas indicate that they differ in mass by 46 mDa. Accurate measurement of the molecular mass for the respective monoisotopic (all ^{12}C, ^{14}N, and ^{16}O) $[M-5H]^{5-}$ species of the complex (m/z 1863.748) and the free RNA (m/z 1784.126) allowed the mass of the ligand to be calculated as 398.110±.009 Da.--

E82

Please replace the paragraph beginning at page 196, line 27 of the specification with the following rewritten paragraph:

--Figure 61 shows high resolution ESI-FTICR spectrum of the library used in Figures 59 and 60, demonstrating that both library members with a nominal molecular weight of 398.1 were present in the synthesized library.--

E83

Please replace the paragraph beginning at page 197, line 3 of the specification with the following rewritten paragraph:

--Based on the high precision mass measurement of the complex, the mass of the binding ligand was determined to be consistent with the library member having a chemical formula of $C_{15}H_{16}N_4O_2F_6$ and a molecular weight of 398.117 Da (Figure 62). Thus, the identity of the binding ligand was unambiguously established.--

E84

Please replace the paragraph beginning at page 197, line 8 of the specification with the following rewritten paragraph:

--Use of exact mass measurements and elemental constraints can be used to determine the elemental composition of an "unknown" binding ligand. General constraints on the type and number of atoms in an unknown molecule, along with a high precision mass measurement, allow determination of a limited list of molecular formulas which are consistent with the measured mass. Referring to Figure 63, the elemental composition is limited to atoms of C, H, N, and O and further constrained by the elemental composition of a "known" moiety of the molecule. Based on these constraints, the enormous number of atomic combinations which result in a molecular weight of 615.2969 ± 0.0006 are reduced to two possibilities. In addition to unambiguously identifying intended library members, this technique allows one skilled in the art to identify unintended synthetic by-products which bind to the molecular target.--

E85

Please replace the paragraph beginning at page 197, line 20 of the specification with the following rewritten paragraph:

--The results of direct determination of solution phase dissociation constants (Kd's) by mass spectrometry is shown in Figure 64. ESI-MS measurements of a solution containing a fixed concentration of RNA at different concentrations of ligand were obtained. By measuring the ratio of bound:unbound RNA at varying ligand concentrations, the Kd was determined by 1/slope of the "titration curve". The MS derived value of 110 nM is in good agreement with previously reported literature value of 200 nM.--

E86

Please replace the paragraph beginning at page 197, line 27 of the specification with the following rewritten paragraph:

--For the determination of ligand binding site by tandem mass spectrometry, a solution containing the molecular target or targets is mixed with a library of ligands and given the opportunity to form noncovalent complexes in solution. These noncovalent complexes are mass analyzed. The noncovalent complexes are subsequently dissociated in the gas phase via IRMPD or CAD. A comparison of the fragment ions formed from dissociation of the complex with the fragment ions formed from dissociation of the free RNA reveals the ligand binding site.--

E87

Please replace the paragraph beginning at page 198, line 7 of the specification with the following rewritten paragraph:

--Figure 65 shows MASS screening of a 27 member library against a 27-mer RNA construct representing the prokaryotic 16S A-site. The inset reveals that a number of compounds formed complexes with the 16S A-site.--

E88

Please replace the paragraph beginning at page 198, line 11 of the specification with the following rewritten paragraph:

--MS/MS of a 27-mer RNA construct representing the prokaryotic 16S A-site containing deoxyadenosine residues at the paromomycin binding site is shown in Figure 66. The top spectrum was acquired by CAD of the [M-5H]5- ion (m/z 1783.6) from uncomplexed RNA and exhibits significant fragmentation at the deoxyadenosine residues. The bottom spectrum was acquired from

E89

E89

by CAD of the [M-5H]5- ion of the 16S-paromomycin complex (m/z 1907.5) under identical activation energy as employed in the top spectrum. No significant fragment ions are observed in the bottom spectrum consistent with protection of the binding site by the ligand.

Please replace the paragraph beginning at page 198, line 19 of the specification with the following rewritten paragraph:

Two combinatorial libraries containing 216 tetraazacyclophanes dissolved in DMSO were mixed with a buffered solution containing 10 mM 16S RNA (see Figure 68) such that each library member was present at 100 nM. The resulting mass spectra, shown in Figure 67 reveal >10 complexes between 16S RNA and library members with the same nominal mass. MS-MS spectra obtained from a mixture of a 27-mer RNA construct representing the prokaryotic 16S A-site containing deoxyadenosine residues at the paromomycin binding and the 216 member combinatorial library. In the top spectrum, ions from the most abundant complex from the first library ([M-5H];5-m/z 1919.0) were isolated and dissociated. Dissociation of this complex generates three fragment ions at m/z 1006.1, 1065.6, and 1162.4 that result from cleavage at each dA residue. More intense signals are observed at m/z 2378.9, 2443.1, and 2483.1. These ions correspond to the w21(3-), a20-B(3-), and a21-B(3-) fragments bound to a library member with a mass of 676.0 ± 0.6 Da. The relative abundances of the fragment ions are similar to the pattern observed for uncomplexed RNA, but the masses of the ions from the lower stem and tetraloop are shifted by complexation with the ligand. This ligand offers little protection of the deoxyadenosine residues, and must bind to the lower stem-loop. The library did not inhibit growth of bacteria. In the bottom spectrum, dissociation of the most abundant complex from a mixture of 16S RNA and the second library having m/z 1934.3 with the same collisional energy yields few fragment ions, the predominant signals arising from intact complex and loss of neutral adenine. The reduced level of cleavage and loss of adenine for this complex is consistent with binding of the ligand at the model A site region as does paromomycin. The second library inhibits transcription/translation at 5 mM, and has an MIC of 2-20 mM against E. coli(imp-) and S. pyogenes.

E90

Please replace the paragraph beginning at page 199, line 13 of the specification with the following rewritten paragraph:

--Figure 68 shows secondary structures of the 27 base RNA models used in this work corresponding to the 18S (eukaryotic) and 16S (prokaryotic) A-sites. The base sequences differ in seven positions (bold), the net mass difference between the two constructs is only 15.011 Da. Mass tags were covalently added to the 5' terminus of the RNA constructs using tradition phosphoramadite coupling chemistry.--

E91

Please replace the paragraph beginning at page 199, line 18 of the specification with the following rewritten paragraph:

--Methodology to increase the separation between the associated signals in the mass spectra was developed in view of the overlap among signals from RNAs 16S and 18S. RNA targets modified with additional uncharged functional groups conjugated to their 5'-termini were synthesized. Such a synthetic modification is referred to herein as a neutral mass tag. The shift in mass, and concomitant m/z, of a mass-tagged macromolecule moves the family of signals produced by the tagged RNA into a resolved region of the mass spectrum. ESI-FTICR spectrum of a mixture of 27-base representations of the 16S A-site with (7 mM) and without (1 mM) an 18 atom neutral mass tag attached to the 5- terminus in the presence of 500 nM paromomycin is shown in Figure 69. The ratio between unbound RNA and the RNA-paromomycin complex was equivalent for the 16S and 16S+tag RNA targets demonstrating that the neutral mass tag does not have an appreciable effect on RNA-ligand binding.--

E92

Please replace the paragraph beginning at page 200, line 3 of the specification with the following rewritten paragraph:

--Paromomycin, lividomycin (MW = 761.354 Da), sisomicin (MW = 447.269 Da), tobramycin (MW = 467.2591 Da), and bekanamycin (MW = 483.254 Da) were obtained from Sigma (St. Louis, MO) and ICN (Costa Mesa, CA) and were dissolved to generate 10 mM stock solutions. 2' methoxy analogs of RNA constructs representing the prokaryotic (16S) rRNA and eukaryotic

E93

E93

(18S) rRNA A-site (Figure 68) were synthesized in house and precipitated twice from 1 M ammonium acetate following deprotection with ammonia (pH 8.5). The mass-tagged constructs contained an 18-atom mass tag ($C_{12}H_{25}O_9$) attached to the 5'-terminus of the RNA oligomer through a phosphodiester linkage.

Please replace the paragraph beginning at page 200, line 26 of the specification with the following rewritten paragraph:

Mass spectrometry experiments were performed in order to detect complex formation between a library containing five aminoglycosides (Sisomicin (Sis), Tobramycin (Tob), Bekanomycin (Bek), Paromomycin (PM), and Livodomycin (LV)) and two RNA targets simultaneously. Signals from the $(M-5H+)^{5-}$ charge states of free 16S and 18S RNAs are detected at m/z 1801.515 and 1868.338, respectively. As shown in Figure 69, the mass spectrometric assay reproduces the known solution binding properties of aminoglycosides to the 16S A site model and an 18S A site model with a neutral mass linker. Consistent with the higher binding affinity of theses aminoglycosides for the 16S A-site relative to the 18S A-site, aminoglycoside complexes are observed only with the 16S rRNA target. Note the absence of 18S-paromomycin and 18S-lividomycin complexes, which would be observed at the m/z's indicated by the arrows. The inset demonstrates the isotopic resolution of the complexes. Using multiple isotope peaks of the $(M-5H+)^{5-}$ and $(M-4H+)^{4-}$ charge states of the free RNA as internal mass standards, the average mass measurement error of the complexes is 2.1 ppm. High affinity complexes were detected between the 16S A site 27mer RNA and paromomycin and lividomycin, respectively. Weaker complexes were observed with sisomycin, tobramycin and bekamycin. No complexes were observed between any of the aminoglycosides and the 18S A site model. Thus, this result validates the mass spectrometric assay for identifying compounds that will bind specifically to the target RNAs. No other type of high throughput assay can provide information on the specificity of binding for a compound to two RNA targets simultaneously. The binding of lividomycin to the 16S A site had been inferred from previous biochemical experiments. The mass spectrometer has been used herein to measure a KD of

E94

E94

28 nM for lividomycin and 110 nM for paromomycin to the 16S A site 27mer. The solution KD for paromomycin has been estimated to be between 180 nM and 300 nM.

Please replace the paragraph beginning at page 202, line 1 of the specification with the following rewritten paragraph:

The 27-mer model of a segment of the bacterial A site region has been prepared as a full ribonucleotide (see Figure 71, compound R), and as a chimeric 2'-O-methylribonucleotide containing three deoxyadenosine residues (see Figure 71, compound C). RNAs R and C have been prepared using conventional phosphoramidite chemistry on solid support. Phosphoramidites were purchased from Glen Research and used as 0.1 M solutions in acetonitrile. RNA R was prepared following the procedure given in Wincott, *et al., Nucl. Acids Res.*, **1995**, *23*, 2677-2684, the disclosure of which is incorporated herein by reference in its entirety. RNA C was prepared using standard coupling cycles, deprotected, and precipitated from 10 M NH_4OAc. The aminoglycoside paromomycin binds to both R and C with kD values of 0.25 and 0.45 micromolar, respectively. The reported kD values are around 0.2 μM. Recht, *et al., J. Mol. Biol.*, **1996**, *262*, 421-436, Wong, *et al., Chem. Biol.*, **1998**, *5*, 397-406, and Wang, *et al., Biochemistry*, **1997**, *36*, 768-779. Paromomycin has been shown previously to bind in the major groove of the 27mer model RNA and induce a conformational change, with contacts to A1408, G1494, and G1491. Miyaguchi, *et al., Nucl. Acids Res.*, **1996**, *24*, 3700-3706; Fourmy, *et al., Science*, **1996**, *274*, 1367-1371; and Fourmy, *et al., J. Mol. Biol.*, **1998**, *277*, 333-345.

E95

Please replace the paragraph beginning at page 202, line 17 of the specification with the following rewritten paragraph:

The mass spectrum obtained from a 5 μM solution of C mixed with 125 nM paromomycin (Figure 72A) contains [M-5H]5- ions from free C at m/z 1783.6 and the [M-5H]5- ions of the paromomycin-C complex at m/z 1907.3. Mass spectrometry experiments have been performed on an LCQ quadrupole ion trap mass spectrometer (Finnigan; San Jose, CA) operating in the negative ionization mode. RNA and ligand were dissolved in a 150 mM ammonium acetate buffer at pH 7.0

E96

with isopropyl alcohol added (1:1 v:v) to assist the desolvation process. Parent ions have been isolated with a 1.5 m/z window, and the AC voltage applied to the end caps was increased until about 70% of the parent ion dissociates. The electrospray needle voltage was adjusted to -3.5 kV, and spray was stabilized with a gas pressure of 50 psi (60:40 N_2:O_2). The capillary interface was heated to a temperature of 180 °C. The He gas pressure in the ion trap was 1 mTorr. In MS-MS experiments, ions within a 1.5 Da window having the desired m/z were selected via resonance ejection and stored with q) 0.2. The excitation RF voltage was applied to the end caps for 30 ms and increased manually to 1.1 Vpp to minimize the intensity of the parent ion and to generate the highest abundance of fragment ions. A total of 128 scans were summed over m/z 700-2700 following trapping for 100 ms. Signals from the [M-4H]4- ions of C and the complex are detected at m/z 2229.8 and 2384.4, respectively. No signals are observed from more highly charged ions as observed for samples denatured with tripropylamine. In analogy with studies of native and denatured proteins, this is consistent with a more compact structure for C and the paromomycin complex. The CAD mass spectrum obtained from the [M-5H]5- ion of C is presented in Figure 72B. Fragment ions are detected at m/z 1005.6 (w6)2-, 1065.8 (a7-B)2-, 1162.6 (w7)2- , 1756.5 (M-Ad)5- , 2108.9 (w21-Ad)3- , 2153.4 (a20-B)3- , 2217.8 (w21)3-, and 2258.3 (a21-B)3- . McLuckey, *et al., J. Am. Soc. Mass Spectrum.*, **1992**, *3*, 60-70 and McLuckey, *et al., J. Am. Chem. Soc.*, **1993**, *115*, 12085-12095. These fragment ions all result from loss of adenine from the three deoxyadenosine nucleotides, followed by cleavage of the 3'-C-O sugar bonds. The CAD mass spectrum for the [M-5H]5- ion of the complex between C and paromomycin obtained with the same activation energy is shown in Figure 72C. No fragment ions are detected from strand cleavage at the deoxyadenosine sites using identical dissociation conditions of Figure 72B. The change in fragmentation pattern observed upon binding of paromomycin is consistent with a change in the local charge distribution, conformation, or mobility of A1492, A1493, and A1408 that precludes collisional activation and dissociation of the nucleotide.

Please replace the paragraph beginning at page 203, line 18 of the specification with the following rewritten paragraph:

--Two combinatorial libraries containing 216 tetraazacyclophanes dissolved in DMSO were mixed with a buffered solution containing 10 μM C such that each library member is present at 100 nM. The resulting mass spectra reveal >10 complexes between C and library members with the same nominal mass. Ions from the most abundant complex from the first library ([M-5H];$^{5-}$ m/z 1919.0) were isolated and dissociated. As shown in Figure 73A, dissociation of this complex generates three fragment ions at m/z 1006.1, 1065.6, and 1162.4 that result from cleavage at each dA residue. More intense signals are observed at m/z 2378.9, 2443,1, and 2483.1. These ions correspond to the $w_{21}^{(3-)}$, a_{20}-$B^{(3-)}$, and a_{21}-$B^{(3-)}$ fragments bound to a library member with a mass of 676.0 = 0.6Da. The relative abundances of the fragment ions are similar to the pattern observed for uncomplexed C, but the masses of the ions from the lower stem and tetraloop are shifted by complexation with the ligand. This ligand offers little protection of the deoxyadenosine residues, and must bind to the lower stem-loop. The libraries have been synthesized from a mixture of charged and aromatic functional groups, and are described as libraries 25 and 23 in: An, *et al., Bioorg. Med. Chem. Lett.*, **1998**, in press. Dissociation of the most abundant complex from a mixture of C and the second library having m/z 1934.3 with the same collisional energy (Figure 73B) yields few fragment ions, the predominant signals arising from intact complex and loss of neutral adenine. The mass of the ligand (753.5 Da) is consistent with six possible compounds in the library having two combinations of functional groups. The reduced level of cleavage and loss of adenine from this complex is consistent with binding of the ligand at the model A site region as does paromomycin. The second library inhibits transcription/translation at 5 μm, and has an MIC of 2-20 μM against E. coli (imp-) and S. pyogenes.--

E97

Please replace the paragraph beginning at page 205, line 17 of the specification with the following rewritten paragraph:

--It is preferred that such data collection and database manipulation be achieved through a general purpose digital computer. An exemplary software program has been created and used to identify the small molecules bound to an RNA target, calculate the binding constant, and write the results to a relational database. The program uses as input a file that lists the elemental formulas of

E98

the RNA and the small molecules which are present in the mixture under study, and their concentrations in the solution. The program first calculates the expected isotopic peak distribution for the most abundant charge state of each possible complex, then opens the raw FTMS results file. The program performs a fast Fourier transform of the raw data, calibrates the mass axis, and integrates the signals in the resulting spectrum such as the exemplary spectrum shown in Figure 74. The peaks in the spectrum are preferably identified via centroiding as shown in Figure 75, are integrated, and preferably stored in a database. An exemplary data file is shown in Figure 76. The expected and observed peaks are correlated, and the integrals converted into binding constants based on the intensity of an internal standard. The compound identity and binding constant data are written to a relational database. This approach allows large amounts of data that are generated by the mass spectrometer to be analyzed without human intervention, which results in a significant savings in time.

Please replace the paragraph beginning at page 206, line 3 of the specification with the following rewritten paragraph:

Figure 74 depicts electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry of a solution which is 5 mM in 16S RNA (Ibis 16628) and 500 nM in the ligand Ibis10019. The raw time-domain dataset is automatically apodized and zerofilled twice prior to Fourier transformation. The spectrum is automatically post-calibrated using multiple isotope peaks of the $(M-5H+)^{5-}$ and $(M-4H+)^{4-}$ charge states of the free RNA as internal mass standards and measuring the m/z difference between the free and bound RNA. The isotope distribution of the free RNA is calculated a priori and the measured distribution is fit to the calculated distribution to ensure that m/z differences are measured between homoisotopic species (e.g. monoisotopic peaks or isotope peaks containing 4 ^{13}C atoms).

Please replace the paragraph beginning at page 206, line 12 of the specification with the following rewritten paragraph: